How effective is a summer school for catch-up attainment in English and maths?

Stephen Gorard, Nadia Siddiqui and Beng Huat See
Durham University
s.a.c.gorard@durham.ac.uk

Abstract

The current study is an efficacy trial of a summer school programme, targeted at literacy and numeracy catch-up for pupils in Years 5 and 6. During the summer holidays the intervention provided extra schooling in English and maths along with a wide variety of extra-curricular activities. The programme developers recruited 435 pupils who volunteered to participate in the programme and agreed to be randomised into treatment and control group. The randomisation resulted in 239 pupils participating in the treatment group and 196 pupils in the control group. The pupils were assessed on Progress in English (PiE) and Progress in maths (PiM) standardised tests after four weeks of the programme. The overall result for English was an 'effect' size of +0.17, with FSM-eligible pupils making even more progress. The effect size in maths was 0, with FSM-eligible pupils making even less progress. As a programme to assist disadvantaged children with their preparation for transition to secondary school, the intervention showed some promise, and most pupils clearly enjoyed their time at the summer school. Parents appreciated the provision of free academic help and enrichment activities over the long summer holiday. But despite considerable efforts from the developers, they recruited fewer than half the intended target of pupils, and attendance once the programme had started reduced the number further. This, added to the cost, suggests that summer schools are not the most effective way to improve attainment for struggling pupils.

1. Introduction

This paper presents the findings of an evaluation of a summer school programme set up by Future Foundations and funded by the Educational Endowment Foundation. It outlines existing evidence on summer schools used to help disadvantaged year 5 and 6 pupils in England improve in maths and English in preparation for their move to secondary school. The paper describes the methods used for the evaluation, the nature and impact of the intervention, and the implications of the study for policy and practice.

In 2012, the Educational Endowment Foundation in England (EEF) supported by Building Educated Leaders for Life (BELL), a major US summer-school provider, funded Future Foundations to run a pilot scheme in Nightingale Academy, Edmonton with the Edmonton Schools Partnership. Funding was provided for a four-week summer school for 160 pupils who had just completed Year 5 or 6. The programme was intended to target pupils from disadvantaged backgrounds, who were underperforming at their expected levels, and likely to

benefit from participation in the programme. The BELL approach has been reported as one of the few forms of summer schools with reasonable evidence of success (Terzian and Moore 2009). One of the reasons for that pilot programme was that there was little robust evidence on the efficacy of the BELL approach in the UK. The pilot was therefore designed to test the feasibility of organising a summer school in a relatively deprived area. In particular, it sought to assess: whether there was demand for the programme, whether families would support and sustain the programme, and whether professional staff would be willing to work during their summer holidays. The report suggested that the approach was feasible, in a suburb of London (Siddiqui et al. 2013), and so a larger trial was set up in more varied locations to establish impact on attainment.

Pupils attending the four-week programme followed a specially designed curriculum involving regular literacy and numeracy lessons taught by trained teachers. Lessons were supported by mentors and peer-mentors and generally conducted in small teaching groups. Each afternoon, students participated in a variety of sports and enrichment activities. The programme took place across three sites in London and the South East: Brighton, Enfield and Islington in the summer of 2013. It was targeted primarily at pupils in Years 5 and 6 who were eligible for free school meals (considered to be disadvantaged by their school) and/or who were not expected to achieve Level 4 in English or maths at the end of Key Stage 2.

2. A summary of the existing evidence

In general, the literature on the effectiveness of summer schools for academic improvement looks promising. A research synthesis of 93 evaluations of summer schools suggested that they can be effective, especially with parental involvement, and perhaps with more promise in maths than literacy (Cooper et al. 2000). Schacter and Jo (2005) found a positive impact for literacy from a seven-week summer school for economically disadvantaged first grade children in the US, randomised to the intervention. Matsudaira (2008) estimated the impact of summer school on both maths and reading as about +0.12, based on a form of regression discontinuity analysis.

However, a problem with all studies requiring voluntary attendance is that those who turn up may be very different from those who do not. For example, Borman and Dowling (2006) compared 438 students randomised to summer school for three years with 248 receiving no intervention. The study reports success for the intervention, but as with almost all such studies the level of dropout is a concern. The study ended up comparing the post-test scores of the ‘compliers’ who continued to attend with all of the control students. This would tend to inflate the perceived impact of the summer school.

Considering preparation by disadvantaged pupils for the transition to secondary school more specifically, a more recent review of literacy interventions identified two studies (See et al. 2012). Neither offered as much promise as the studies outlined above. Sainsbury et al. (1998) randomly allocated around 2,000 pupils to either a literacy summer school or a control. Both groups had worse scores at the end of the summer than at the beginning. Kim (2006) randomly allocated 331 pupils from grades 1 to 5 (and therefore few in the age group

relevant to this paper) to either a free book treatment group or a waiting-list control. Each pupil received 10 free books over the summer, and materials and support for reading such as letters and postcards. The treatment group reported reading three more books on average than the control. However, there was no difference in their literacy scores on return to school. Neither study means that a summer school intervention cannot work. But they do suggest that the decline in literacy over the summer of transition to senior school is not simply about lack of literacy practice or activity.

There have been several evaluations of the BELL summer schools in the US more specifically (BELL 2001, 2002, 2003). These follow the format adapted by Future Foundations for this new intervention in England. Unfortunately, their own 'evaluations' are often unclear, and reinforce the importance of independent evaluation where concern is more about finding the impact than in what that impact is. The BELL programme took place in several US cities, but the reports are not always clear which sites provided data for which analysis, nor how the students were allocated to the analytical groups. There appears to have been no benefit for writing (Harvard Family Research Project 2006). Gains are reported but no effect sizes were published. The gains were lower for low-income children and those in the age range relevant to school transition.

A similar study by researchers with a potential conflict of interest looked at BELL summer schools in two US cities (Chaplin and Capizzano 2006). Applicants were randomised to receive the 'treatment' or not. The treatment involved tuition in maths and reading but no maths scores are presented in the report. The overall 'effect' size for reading calculated here (not by the original authors) was +0.02 - which is negligible. More importantly, 46% of those randomised dropped out or refused to continue with the study, and the results are available for only 44% of the initial randomised students. Overall therefore, despite some claims to the contrary, there is no strong evidence that the BELL approach would work in England with disadvantaged pupils preparing for secondary school.

In September 2011, the Deputy Prime Minister announced that £50 million would be made available for a summer schools programme in England. The scheme was intended to support disadvantaged pupils in the transition phase from primary to secondary school. Over 1,700 schools conducted summer school programmes in 2012, and in the summer of 2013 over 1,900 schools conducted summer programmes sponsored by the Department for Education. An evaluation of summer school impact involved around 21,000 pupils in schools that conducted the summer school programme and pupils in comparator schools that did not participate (Marting et al. 2013). The 'evaluation' did not look at the impact on attainment, and was solely concerned with survey responses to items about pupils' confidence, social skill and readiness to attend their secondary school.

The EEF has commissioned two other randomised controlled trials to assess the impact of summer schools on pupils' performance in reading and writing (Torgerson et al. 2014, Maxwell et al. 2014). In both studies common challenges were the initial recruitment of pupils to the summer schools and pupil drop-out during the programme. Improvement in reading comprehension was reported by both studies. This new study is larger and looks at both English and maths, as well as involving a full process evaluation. The intervention itself

is described in part 5 of the paper, as part of the process evaluation which contrasts, where necessary, what was planned and what was observed.

3. Methods of evaluation

3.1 Trial Design

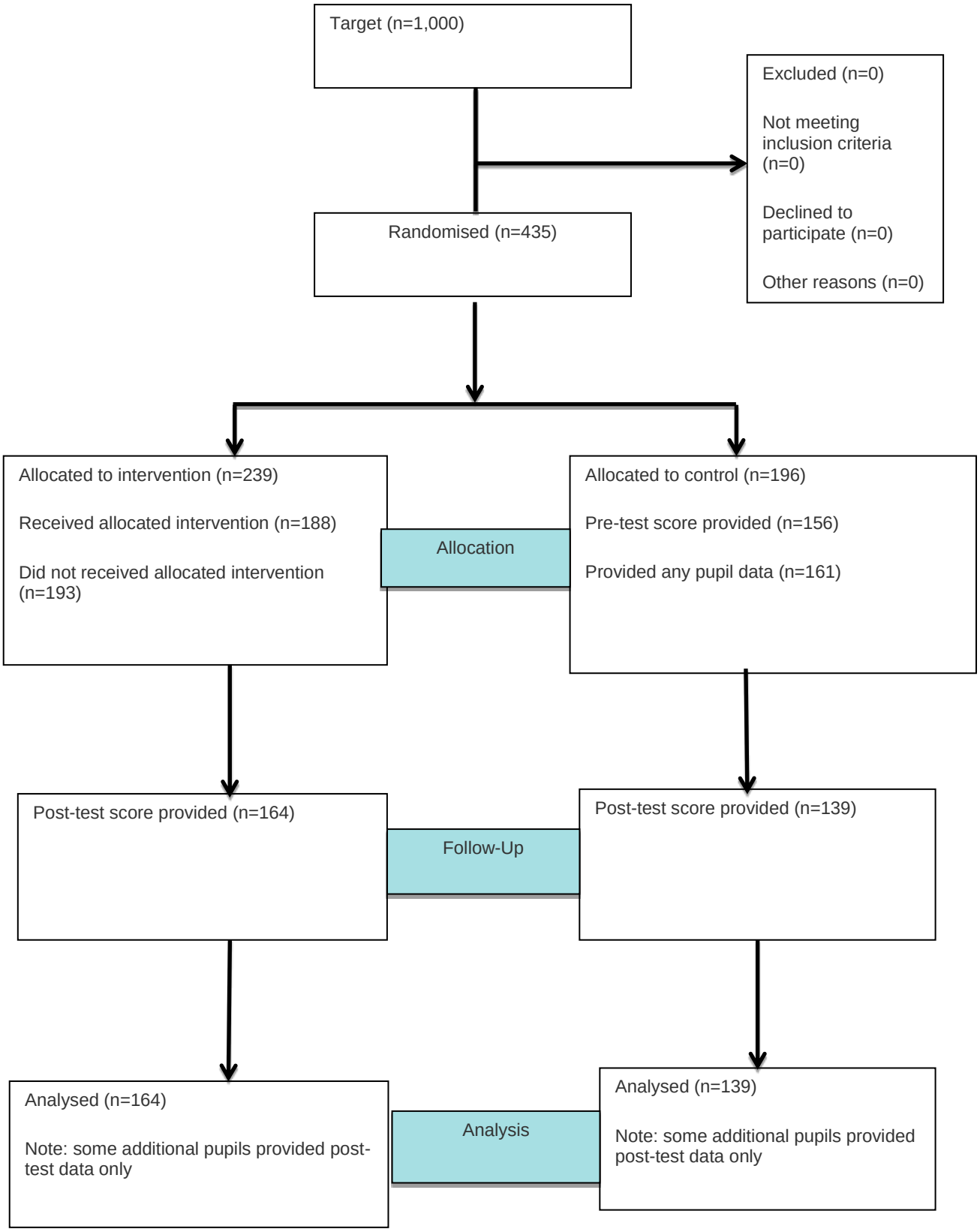
The intervention was proposed and designed as a simple individually randomised controlled trial of 1,000 pupils divided into two equal groups – a treatment group and a control without placebo or waiting list. The first group received the treatment over summer 2013. The prior attainment scores consisted of KS2 fine point scores from summer term 2013 (including teacher assessed scores for year 5 pupils). These were obtained from schools via the developer and represent each child's summer test scores for reading, writing and maths. Post-tests (GL Assessment's Progress in English and Progress in Maths) were administered in groups in the eventual schools attended by both groups of pupils in autumn 2013. Primary schools were generally co-operative in conducting the pencil and paper post-tests delivered to each of the schools by the GL assessment team. It was harder to get agreement from secondary schools to test the original Year 6 pupils after they had begun Year 7.

A great deal of effort was put in to reduce demoralisation and consequent dropout. This involved not revealing the groups until after the randomisation, use of a refundable deposit for registrants, and neutral administration of the post-test.

3.2 Sample

The target sample was expected to be 1,000 pupils for years 5 and 6. However, the developers found recruitment much harder than anticipated. One of the original sites was abandoned due to the partner selected to host not being able to secure a school committed to the full trial. Located in the most deprived regions, 43 schools referred the programme to parents and carers whose children needed the catch-up intervention in their schools. In total, only 435 individual pupils volunteered to take part in the study, hoping to be part of the summer programme (Figure 1). Of these, only 303 provide full matched data. Attrition was, therefore, 30% of the original sample.

Figure 1 – Participant flow diagram


Target (n=1,000)
Excluded (n=0)
Not meeting inclusion criteria (n=0)
Declined to participate (n=0)
Other reasons (n=0)
Randomised (n=435)
Allocated to intervention (n=239)
Received allocated intervention (n=188)
Did not received allocated intervention (n=193)
Allocation
Allocated to control (n=196)
Pre-test score provided (n=156)
Provided any pupil data (n=161)
Post-test score provided (n=164)
Follow-Up
Post-test score provided (n=139)
Analysed (n=164)
Note: some additional pupils provided post-test data only
Analysis
Analysed (n=139)
Note: some additional pupils provided post-test data only

3.3 Randomisation

The randomisation was carried out in three waves so that each batch of applicants would know early whether they were free for summer or not. The evaluators also agreed with the developers and funders to have a larger treatment than control group to make the programmes viable. In addition, funders requested that more Year 6 pupils be assigned to the intervention group. All of these factors mean that the resultant allocation was not a simple randomisation process (even before attrition).

A mechanically-shuffled set of cards was used for generating random numbers. There was one card for each pupil, with odd/even representing the two groups. Each pupil was allocated to a group according to the card dealt for them at their turn on the list. The list with this group allocation was then sent back to developers to inform the schools and parents for consent and to register the relevant pupils to attend the summer school programme. One pair of twins was treated as one case at the request of the developers meaning that in effect there were 434 cases.

The randomisation resulted in the allocation of 239 pupils to the treatment group to attend the summer schools programme, and 196 pupils to the control which means they were simply followed for the post-test. In the final analysis the sample retained had 303 pupils – with 75 treatment and 30 control pupils not included in the final analysis. The former were largely those allocated to treatment but then not turning up to the summer school. The reasons given were that parents had work or holiday arrangements that clashed, or pupils were ill, changed their minds, or did not want to attend as their friend(s) had not been selected. The missing control pupils were largely those whose secondary school would not conduct the test.

The pupil characteristics in the achieved sample are provided in Table 1. The developers were successful in targeting pupils with various possible indicators of disadvantage, and these characteristics were proportionately very similar for both achieved groups.

Table 1 – Percentage of pupils with specific background characteristics, by treatment group

Category	Treatment	Control
Original Year 5	68	66
Boys	52	58
FSM-eligible	44	48
Having English as an additional language	36	36
SEN (statemented, School Action or Action plus	31	32
Ethnic minority (other than White UK)	69	74
Total	193	161

Note: Year 5 pupils are the 9-year-old cohort. Year 6 are 10-year-olds.

3.4 Analysis

The trial outcomes were assessed by means of the GL Assessment Progress in English (PiE) and Progress in Maths (PiM) test given to all relevant pupils in the term following the summer school. All pupils originally identified as eligible were tested and their outcomes analysed, as far as possible, regardless of the time actually spent on the intervention (intention to treat).

The difference between the two groups in raw score terms, contextualised by pupils' prior KS2 fine point scores in reading, writing and maths, was converted into an 'effect' size of the gain scores (Hedges' g). 'Gain' scores refer to the average difference in scores between the prior attainment in KS1 fine point scores and the post-test raw scores. This approach was considered necessary here because the groups were unbalanced at the outset due to the high level of dropout after individual randomisation. It would have been preferable to have had pre-test scores from exactly comparable standardised tests to the post-tests. However, the GL Progress tests selected by the developer does not have comparable alternative versions, and the funders wished to reduce the cost of what had been planned as a 1,000 case trial anyway. The pre-test scores were re-scaled before further analysis to match the range of the post-test raw scores.

A further analysis included using the same 'effect' size for FSM-eligible pupils only.

Originally, the evaluation was intended to include an analysis of dosage. However, the attendance figures provided by the developers were not complete enough to conduct this.

It is possible that some readers will want to understand why more complex analyses, involving significance testing and similar, were not conducted here. There are four main over-lapping reasons. The design work undertaken for an RCT pre-specifies much of the analysis, as opposed to studies with weaker designs that rely on *post hoc* discovery (Gorard 2013). There are two (or more) groups in an RCT, and the difference between them in terms of outcome measure(s) is the result. Approaches based on standard errors, such as significance tests, do not answer the substantive question that an RCT is addressing. If they worked, significance tests would only help assess whether the result was due to the vagaries of random allocation. However, they do not work as intended (Carver 1978), are widely misunderstood (Watts 1991), and can produce dangerously misleading results (Falk and Greenbaum 1995). Finally, and perhaps most simply, they would not be relevant here anyway as the 30% dropout after randomisation means that the two groups are no longer random, and so the question of whether the result was due to the vagaries of random allocation has no meaning (Gorard 2015). These approaches are necessarily predicated on random sampling/allocation, and take no account of sample quality or attrition (Berk and Freedman 2001). Effect sizes are the preferred approach (Lipsey et al. 2012), and are used here.

3.5 Process evaluation

The process evaluation was conducted by the independent evaluators, with the full co-operation of the developers. The latter conducted the training of staff, monitored the intervention, and collected formal records and the views of staff. The evaluators observed the

training, the teaching and testing, used the texts and documents relating to the intervention, and conducted face-to-face interviews with staff, students, parents and project members. These interviews were usually conducted without a formal pre-set schedule, and arose as the situation allowed. The observations of staff training and implementation of the programme in action were as simple and integrated and non-intrusive as possible.

Attention was given to the training of staff, teachers, mentors and peer-mentors, their receptivity to the programme and delivery of the intervention at the three different sites. Two evaluators attended the first day of the training when academic content was delivered to the selected teachers. The evaluators were participant observers, as which they talked to teachers, mentors and peer-mentors, teacher trainers and project leaders to get their perceptions of the programme and the process of curriculum development teachers' recruitment and teachers' selection criteria.

The teachers selected for the summer school were all well-qualified and experienced in teaching at primary or secondary levels. The teacher selection criteria were decided by the developers' team aiming to have an effective model of teaching for the four-week programme. The subsequent fieldwork in schools included observations of these trained staff delivering the intervention, noting inconsistencies or any departures from the programme protocol. The evaluators also considered the resources used (such as portfolios, reading materials, feedback sheets), and asked for staff, mentors' and peer-mentors' and pupils' perceptions of these materials. The interviews and field notes were part transcribed and shared between the evaluation team. Any school and pupil names have been altered to pseudonyms.

The process evaluation was useful in assessing fidelity to treatment. The perceptions of participants provided indications of any resentment or resistance to the programme. They are also useful in identifying potential issues or barriers which could be addressed for any future scaling up.

4. What difference did it make to attainment?

The two treatment groups were slightly unbalanced in terms of the available prior performance from the outset. The treatment group is ahead in terms of KS assessment scores in reading, writing and maths (Tables 2 to 4). This means that a straightforward post-test only analysis would be misleading. The question is therefore whether the treatment group has improved from this position of relative advantage at the outset.

Table 2 - KS2 Reading score, pre-test, all achieved scores

	N	Mean	Standard deviation	'Effect' size
Treatment	169	23.66	4.86	+0.02
Control	144	23.42	5.76	-

	N	Mean	Standard deviation	'Effect' size
Overall	313	23.55	5.29	-

Note: several pupils with post-test scores had no prior data provided by schools.

Table 3- KS2 Writing score, pre-test, all achieved scores

	N	Mean	Standard deviation	'Effect' size
Treatment	169	22.57	5.03	+0.11
Control	144	21.94	5.97	-
Overall	313	22.28	5.48	-

Table 4 - KS2 Maths score, pre-test, all achieved scores

	N	Mean	Standard deviation	'Effect' size
Treatment	169	23.32	6.40	+0.06
Control	144	22.89	8.12	-
Overall	313	23.12	7.23	-

The treatment group also has higher post-test scores on both outcomes, assessed by Progress in English and Progress in Maths (Tables 5 and 6). In English the relative advantage for the treatment group is higher than in the pre-test scores, suggesting a small gain of around +0.06 (+0.17-0.11) to +0.15 (0.17-0.02). In maths the relative advantage for the treatment group is lower than in the pre-test scores, suggesting a small loss of around -0.02 (+0.03-0.05). Overall, therefore, and coupled with the problem of attrition, this is not particularly secure evidence of a benefit for attainment from attending a summer school.

Table 5 - PiE score, post-test all achieved scores

	N	Mean	Standard deviation	'Effect' size
Treatment	191	39.06	16.08	+0.17
Control	160	36.36	15.41	-
Overall	351	37.83	15.82	

Note: several pupils with post-test scores had no prior data provided by schools.

Table 6 - PiM score, post-test all achieved scores

	N	Mean	Standard deviation	'Effect' size
Treatment	187	20.87	10.07	+0.05
Control	159	20.39	11.02	-
Overall	346	20.65	10.50	

Note: the same result is found when the calculations ignore all cases with any missing pre-test scores, and the overall N drops to 313, as in Tables 2 to 4.

These results are confirmed when treating the rescaled KS2 fine points scores as if they were true pre-test data, and calculating 'gain' scores as the difference between PiE or PiM and the pre-test. The pre-test score for English is calculated as the mean of the scores for reading and writing. When the 'gain' scores from rescaled KS fine points to PiE or PiM raw scores are considered, there is again a positive result for English and a very small negative one for maths (Tables 7 and 8). It is noticeable that, insofar as the scores mean anything comparable, the gain in maths over the summer was considerably greater than in English for both treatment and control. It is possible either that maths is consolidated over summer more than English even without practice or otherwise more resilient to summer learning loss. English skills, on the other hand, may need the practice that the summer school is one way of providing.

Table 7 - Gain score English, all achieved scores

	N	Mean	Standard deviation	'Effect' size
Treatment	167	3.71	13.24	+0.17
Control	143	1.46	12.79	-
Overall	310	2.67	13.06	-

Table 8 - Gain score Maths, all achieved scores

	N	Mean	Standard deviation	'Effect' size
Treatment	164	7.51	8.43	-0.00
Control	142	7.53	9.24	-
Overall	306	7.52	8.80	-

Considering only the achieved sample for FSM-eligible pupils, the results are similar to the overall results.

5. What was the summer school programme like?

Over the four weeks, all pupils participated in two 75 minute academic lessons each morning, one for literacy and one for numeracy. These were delivered to small tuition groups of a planned 10+ pupils by high-quality teachers, supported by two mentors (one of which was a sixth-former or other student), using a scheme of work designed specifically for the programme. In the afternoons, pupils took part in a diverse set of enrichment activities, including sports, arts, cookery, off-site theatre and raft building. The programme closed with a special Graduation ceremony for pupils and parents/carers.

5.1 Academic sessions

Each day included small group teaching in literacy and numeracy. The main aspects of the academic activities were:

- Structured lesson plans (75 minutes for each class session for numeracy and literacy) with clearly written aims for every activity.
- A specific summer programme reading book
- Individual pupils' portfolios
- Teacher evaluation and feedback

5.2 Lesson plans

The lesson plans were designed by experts in each area in preparation for the summer, using formative feedback from the pilot study. Teachers were given details on the aims of each lesson, and the activities and resources they were supposed to conduct in the sessions
An example of each lesson was observed by one lead coordinator in all three sites. Their job was to give feedback to the teachers, and by regularising the process of teacher observation and feedback to standardise the intervention all across the sites and maintain quality of teaching. This feedback model was similar to the Ofsted model for teacher evaluation.

5.3 Pupils' portfolio

Pupils were given portfolios which were complete activity books for literacy and numeracy learning activities. These included the resources, worksheets, evaluations and progress reviews for every lesson of the 4 weeks. They were devised to allow gradual progress in performance, and demonstrate this to the children, their parents and those evaluating the programme. Each portfolio was also a record of pupils' learning.

5.4 Teacher feedback to students

All students were seen working on the portfolio and teachers/mentors were giving their feedback on performance. Other regular rewards included stars, stickers, and merit awards.

5.5 Student feedback to teachers and mentors

At the end of each lesson there was a student feedback corner to tell the teacher whether they had achieved their learning objectives, enjoyed the activity, or found difficulties.

5.6 Teacher transition in Week 3

The teachers were generally recruited to work for only two of the four weeks. In the third week there was a teacher's transition phase and a new batch of teachers took over. A handover session was organised on the Friday of week 2 to enable all new teachers to meet their colleagues, learn about pupils and develop an understanding of the rhythms and routines on their site. The mentors and peer mentors remained the same. According to the developers this transition was necessary in order to get good teachers to work during the summer holidays and to ensure they would be able to give them full energy to the programme and return to school rested. They could only engage many working teachers for no more than two weeks so in order to make it standard practice all teachers (even if they could work additional weeks) changed in the third week. Teachers in both of the phases were given the same training and were informed about this change when they were first recruited. The portfolio was a record of each pupil performance so that the second batch of teachers could assess how each pupil had been performing previously.

5.7 Progression Monitoring Sheet (PMS)

The PMS was an addition to the programme to help ensure the smooth transition between teachers. This was completed by mentors working closely with pupils, and made available to the teachers. Feedback from mentors who were likely to be around for the whole duration of the programme was found useful by the new group of teachers starting in the 3rd week.

5.8 Enrichment activities

The afternoon sessions were for a range of enrichment activities. The dvelopers hired external providers according to each site who could arrange a variety of activities over the 4 weeks. The activities included drama performance, art and creativity, science challenges, cooking, dance, music, and sports such as football, indoor cricket, fencing, and gym exercises. There were also arrangements for outdoor activities such as raft building, swimming and theatre performance. To some extent these activities varied by site with London having a greater range of cultural and other resources within easy travel distance.

The staff training was conducted for all teachers, mentors and peer-mentors who were planned to be employed at three different sites. In the event not all could be employed due to low recruitment of pupils. The training was well conducted and structured, with many examples of activities which teachers could use. Teachers were generally motivated and eager to try something new. They seemed positive and to enjoy the activities. The role of mentors was defined in terms of academic support to pupils in the classroom rather than just assisting teachers in executing the lessons. During the visits, mentors were seen supporting pupils in the classrooms, and specifically those who regularly needed the individualised attention of an adult.

A primary school model was followed for the settings in which pupils do not have to change the classrooms. Teachers came to the classrooms and all teaching materials are made available to the pupils there. This model was adopted with a view that primary school pupils are more used to this format of classes and there would be more chance to emphasise teaching and learning activities rather than making them adjust to the new routines. On the other hand, this reduced the chances for experiencing what would happen in secondary school.

The planned literacy and numeracy sessions were closely followed on all sites. The implementation of literacy sessions was very close to the developed lesson plans. However, numeracy sessions were observed to change with time into individualised tutoring sessions. One of the reasons was perhaps a wide range of abilities in some class groups, and teachers with the help of mentors broke numeracy classes into further small groups or one to one sessions.

The number of pupils in each classroom was fewer than expected. There were usually 3 or even 4 adults in each classroom of 10+ pupils, and every child was seen receiving extra attention and time from the adults. In some observations there were more adults than pupils in each classroom. This was partly due to low recruitment, and partly due to pupil absences. Once class observed was at half strength. The results of the current programme need to be interpreted in the light of these high ratios.

The intervention project team had a system to monitor and feedback to teachers on all three sites. Each school had been assigned a coordinator who looked at progress on the syllabus followed in the teaching, and gave regular feedback to teachers on the quality of teaching. In one case it was reported that a teacher was asked to leave the programme following multiple unsatisfactory observations.

Several lessons observed by the evaluators were poorly taught especially for maths. Basic pedagogical and factual errors were observed, and in one case pupil written responses were marked incorrectly. In literacy especially, more sessions were seen to be fun and enjoyable for all. Despite the pupil ratios, class control was sometimes poor.

5.9 The cost of the programme

The achieved pupil to teacher ratio was around 5.5, and each teacher class also had a peer-mentor and a mentor was available. The total cost of provision was between £500,000 and £550,000 for 193 pupils (not all of whom participated in this evaluation) across three sites. Therefore, the programme cost something like £2,800 per pupil. This is therefore an expensive intervention when considered only in terms of improving attainment.

5.10 Barriers to delivery

All teaching staff were committed to two weeks of the summer school programme. In the third week there was a change in most teaching staff which was a major transition to adjust to in a short programme of just 4 weeks. In order to overcome the effect of this transition mentors and peer-mentors remained the same. The mentors kept regular monitoring details of

each class they were supporting. This monitoring record was given to the new teachers in the 3rd week in order to help them understand individual pupils' academic performance and behaviour in the class. This seemed to work better than it had in the pilot.

The classes were combined year groups. The older pupils were sometimes more aware of the concepts being taught in the class and therefore the given tasks were easier for them to complete. It was also observed that Year 6 pupils were doing some activities for the Year 5 pupils, and so not letting them learn independently. The groups were not homogeneous in terms of abilities and levels. Some pupils were not necessarily in need of a literacy and numeracy catch-up intervention, and the taught content was not up to the level of all pupils, especially those who were in year 6. This variety was sometimes a challenge for teachers and perhaps the reason that teachers were observed losing control (even in classes with fewer than two pupils per adult). In some of the observed lessons, more time was given to class management instructions than the teaching task or discussions on the topic.

The developers recruited teachers according to the criteria of teaching experience, enthusiasm for teaching and willingness to support disadvantaged pupils. However, even so the standard of teaching was mixed.

5.11 Is the intervention attractive to stakeholders?

This summer school programme was created as a literacy and maths catch-up intervention to take place during the summer holidays. The curriculum and syllabus is not much different from what is followed in ordinary schools, which some of the pupils most in need of catch-up may not enjoy (although this was not a view expressed particularly strongly by the pupils themselves). Extra activities were integrated which supported the learning. Overall, the atmosphere in each site was good.

Interviews with parents suggested that the free summer programme kept their children engaged in reading and maths. Parents reported their satisfaction concerning the quality and overall management of the programme. Some of them also stated that they could not afford expensive holiday trips for their children so such an arrangement provided opportunities to their children which they could not have managed otherwise.

Pupils were generally enthusiastic about attending the summer schools. Some of them also reported that it was the afternoon activities they enjoyed more than the teaching sessions. They said that the teachers and mentors were very supportive and several pupils wanted to come back to the programme in the next summer.

5.12 Perceived outcomes

Teachers in the summer schools reported that the syllabus given to them was very well structured. They did not need to prepare more than was given in the detailed lesson plans, although some further planning and adaptation was possible. The objectives, timings, use of materials and feedback, according to teachers, were all fully integrated in the plans. However, the teachers reported that this was not always appropriate when dealing with groups of mixed

year groups and abilities. Some flexibility was needed. The teachers for numeracy were also observed to have divided the numeracy classes in smaller groups of 2 to 3 pupils as they found that not all of the pupils could follow similar plans. In one numeracy session pupils with SEN were working with mentors all of the time and they were given materials and practice sheets which were different from the syllabus planned for the whole class. This was less an issue for literacy activities.

The mentors working with pupils with special educational needs reported that regular tutoring support had increased the confidence of pupils and in the 4th week of the programme they felt that the pupils had improved. Pupils on all three sites were observed to have received support in literacy and numeracy almost at the level of private tutoring. The teachers suggested that individualised attention had increased the confidence of pupils.

Regular attendance of summer school was also perceived to be an indicator of success by parents, teachers and mentors. Most pupils, on all three sites, were reportedly attending the school regularly, and they had enjoyed coming despite it being during the holidays. Because the programme was in a secondary school, pupils could gain a different type of school-experience, and a different approach to teaching and learning. This could help them with the transition to secondary school. Also, through being out of their usual roles as possible 'naughty' or 'silly' kids they could have a fresh start with all new teaching and management staff members. Mixing the year groups 5 and 6 was perceived by some pupils as a fun experience and a positive influence on their learning.

5.13 Formative findings

The programme was intended to support struggling pupils to catch-up in literacy and numeracy learning. The pupil selection and recruitment procedure was mainly dependent on volunteering pupils and parents, and recruitment was difficult. Whilst recruitment materials and messaging for schools made clear the pupil selection criteria, limited access to pupil demographic data early enough before the intervention meant it was impossible to verify all applications. Therefore the final selection had a higher proportion than desirable of pupils who were above the intended levels of the programme. This needs to be prevented in any future activities.

Similarly, the pupils were divided into class groupings on the basis of little information about their needs for literacy and/or numeracy. In order to manage poor behaviour and discipline some groupings were changed, but apparently no such decision was taken on the basis of pupils' performance. The short span of four weeks is very limited to identify and make changes, but a prior consideration of the needs of the selected pupils might have improved the effectiveness of the targeted programme.
The evaluators observed that pupils with behavioural issues were taken out of the sessions for mentoring. They met with a behaviour mentor to discuss their behaviour and followed a process for being reintroduced to the class. This was introduced in the 2012 pilot and specifically trained staff from secondary schools were recruited and selected for this role. This was done both to help prepare that pupil for returning to class, and to maintain the flow and

discipline of the classes, so avoiding the loss of learning for other pupils. However, there was no extra academic hour or activity to support that child's loss in learning.
The post-tests used were the pencil-and-paper versions of PiE and PiM. GL Assessment developed these standardised tests which were posted to each of the addresses of the schools attended by pupils who had participated in the programme. The evaluation team members visited some schools to observe the testing process, which was found to be good. Some of the SEN pupils needed staff support to read the questions to them and this was observed to be done in a very neutral way. Most of the secondary schools now attended by the previous year 6 pupils were not previously part of the intervention, and some were less co-operative with the post-test than the primary schools. This led to higher attrition of scores for year 6 than year 5.

6. Conclusions

6.1 Limitations of study

The findings are based on a randomised controlled trial, with individual random allocation to groups and a waiting list for pupils who were initially not selected to receive the intervention. The evaluation was set up as an efficacy trial to test the impact of the summer school in the best possible conditions to see if they hold promise, but did not seek to demonstrate that the findings hold at scale in all types of schools. To test this question, a future evaluation could be run on a larger scale in a wider variety of areas. It is not clear from this short-term evaluation what the longer-term benefits could be.

The study was designed to involve 1,000 students. However, the problems with attracting students and keeping them in the project mean that the results are substantially weakened. The actual population of students who fitted the criteria set out for the programme caused an issue, and were exhausted – targeted recruitment took place across a wide range of schools. The ability to make any generalised comments about this summer programme is limited, other than the difficulty in recruiting also encountered by other studies, because of the nature of the attained sample, and high attrition after randomisation.

6.2 Interpretation

The pre-existing evidence on the impact of summer schools is very mixed. A number of studies have been conducted, most commonly in the US, but these have often been methodologically weak, meaning that further study in this area is required.

However, though caution is essential, there is some promise from the results here. This was a well-run intervention, generally supported by teaching staff and pupils. To achieve this level of co-ordination and support required considerable energy and monitoring. The class sizes were very small, making provision even more expensive than planned. If the primary purpose is to enhance literacy and maths catch-up there are already exist more effective and cheaper alternatives. Indeed, the evidence is that this programme provides no benefit to attainment in maths. This could have been due to a difference in the quality of the teaching observed, or it

could be that maths is less susceptible than literacy to summer learning loss. There is more promise for attainment in literacy, but it is no more than a strong promise until demonstrated more convincingly (perhaps using a cross-control design that commits all volunteers to some summer activity whether summer school or an unrelated alternative). The probable gain in literacy is no more than the number of weeks of actual literacy teaching.

The programme was implemented successfully on all three sites. Despite considerable efforts from the developers, a significantly smaller number of pupils attended the school than had been hoped for, with less than half of the target number of students signing up for the programme. These challenges were particularly apparent in the numbers recruited outside of London (only 19% of the target in Brighton compared to 34% of target in Islington and 44% in Edmonton), suggesting that there may be problems were the programme to be rolled out, particularly to less densely populated areas.

It is clear that many pupils enjoyed their time at the summer school, and it is likely that the programme led to non-academic benefits for at least some of their families. The programme was popular with parents, who appreciated the free provision of academic and enrichment activities over the summer holidays.

Acknowledgements

This evaluation was funded by the Educational Endowment Foundation.

References

BELL (2001) *BELL Accelerated Learning Summer Program 2001 evaluation report*, Dorchester, MA

BELL (2002) *BELL Accelerated Learning Summer Program 2002 national evaluation report*, Dorchester, MA

BELL (2003) *BELL Accelerated Learning Summer Program: 2003 program outcomes*, Dorchester, MA

Berk, R. and Freedman, D. (2001) *Statistical assumptions as empirical commitments*, http://www.stat.berkeley.edu/~census/berk2.pdf, accessed 030714

Borman, G. and Dowling, N. (2006) Longitudinal Achievement Effects of Multiyear Summer School: Evidence From the Teach Baltimore Randomized Field Trial, *Educational Evaluation and Policy Analysis*, 28, 1, 25-48

Carver, R. (1978) The case against statistical significance testing, *Harvard Educational Review*, 48, 378-399

Chaplin, D. and Capizzano, J. (2006) *Impacts of a summer learning program: a random assignment study of Building Education Leaders for Life (BELL)*, Washington, DC: The Urban Institute, http://www.urban.org/UploadedPDF/411350_bell_impacts.pdf

Cooper, H., Charlton, K., Valentine, J. and Muhlenbruck, L. (2000) *Monographs of the Society for Research into Child Development*, 65, 1

Falk, R. and Greenbaum, C. (1995) Significance tests die hard: the amazing persistence of a probabilistic misconception, *Theory and Psychology*, 5, 75-98

Gorard, S. (2013) *Research Design: Robust approaches for the social sciences*, London: SAGE

Gorard, S. (2015) Rethinking "quantitative" methods and the development of new researchers, *Review of Education*, 3, 1, 72-96, doi: 10.1002/rev3.3041

Harvard Family Research Project (2006) *Evaluation of the BELL (Building Educated Leaders for Life) Accelerated Learning Summer Program*, http://www.hfrp.org/out-of-school-time/ost-database-bibliography/database/bell-accelerated-learning-summer-program/evaluation-1-2002-national-evaluation-report

Kim, J. (2006) Effects of a Voluntary Summer Reading Intervention on Reading Achievement: Results from a randomized field trial, *Educational Evaluation and Policy Analysis*, 28, 4, 335-355

Lipsey, M., Puzio, K., Yun, C., Hebert, M., Steinka-Fry, K., Cole, M., Roberts, M., Anthony, K. and Busick, M. (2012) *Translating the statistical representation of the effects of education interventions into more readily interpretable* forms, Washington DC: Institute of Education Sciences

Marting, K, Sharp, C. and Mehta, P. (2013) *The impact of the summer schools programme on pupils*, Report to National Foundation for Educational Research, https://educationendowmentfoundation.org.uk/uploads/pdf/EEF_Evaluation_Report_-_Summer_Active_Reading_Programme_-_October_2014.pdf

Matsudaira, J. (2008) Mandatory summer school and student achievement, *Journal of Econometrics*, 142, 2, 829-850

Maxwell, B., Connolly, P., Demack, S., O'Hare, L., Stevens, A., and Clague, L. (2014) *Summer active reading programme: Evaluation report and Executive.* London: Educational Endowment Foundation.

Sainsbury, M., Whetton, C., Keith, M. and Schagen, I. (1998) Fallback in attainment on transfer at age 11: evidence from the Summer Literacy Schools evaluation, *Educational Research*, 40, 1, 73-81

Schacter, J. and Jo, B. (2005) Learning when school is not in session: a reading summer day-camp intervention to improve the achievement of exiting First-Grade students who are economically disadvantaged, *Journal of Research in Reading*, 28, 2, 158–169

See, BH and Gorard, S. (2014) Improving literacy in the transition period: a review of the existing evidence on what works, *British Journal of Education, Society and Behavioural Sciences*, 4, 6, 739-754, http://www.sciencedomain.org/issue.php?iid=431&id=21

Siddiqui, N., Gorard, S. and See, BH (2014) Is a summer school programme a promising intervention in preparation for transition from primary to secondary school?, *International Education Studies*, 7, 7, 125-135

Terzian, M. and Moore, K. (2009) *What works for summer learning programs for low-income children and youth?*, Washington: Child Trends, http://www.wallacefoundation.org/knowledge-center/summer-and-extended-learning-time/summer-learning/Documents/Effective-and-Promising-Summer-Learning-Programs-Fact-Sheet.pdf

Torgerson, C., Torgerson, D., Jefferson, L., Buckley, H., Ainsworth, H., Heaps, C. and Mitchell, N. (2014) *Discover summer school: Evaluation report and Executive summary*. London: Educational Endowment Foundation. 2-25,

https://educationendowmentfoundation.org.uk/uploads/pdf/EEF_Evaluation_Report_-_Discover_Summer_School_-_May_2014.pdf

Watts, D. (1991) Why is introductory statistics difficult to learn?, *The American Statistician*, 45, 4, 290-291